[redacted]

[redacted]

21 p.

N64 13284
CODE-1

Technical Report No. 32-524 (NASA CR-53016; JPL-TR-32-524)

Continuous Estimation of Sequentially Correlated Random Variables

C. G. Pfeiffer

OTS: $2.60 ph
$0.83 mf

OTS PRICE

XEROX $ 2.60 ph.

MICROFILM $ 0.83 mf.

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

October 30, 1963 21 p refs

Technical Report No. 32-524

Continuous Estimation of Sequentially Correlated Random Variables

C. G. Pfeiffer

T. W. Hamilton

T. W. Hamilton, *Chief*
Systems Analysis

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

October 30, 1963

CONTENTS

ABSTRACT

13284

The concept of a stochastic process is discussed, and sequential correlation is defined. Equations are presented for obtaining the minimum variance estimate of the state of the process when given linearly related data sampled at discrete times. These results are extended to the continuous case, yielding a linear differential equation for the time-varying estimate. This equation is solved to obtain an integral representation of the estimate. An application of the estimation technique is developed.

AUTHOR

I. INTRODUCTION[1]

This paper is concerned with the task of continuously estimating the time-varying state of a stochastic process when given data which are linearly related to that state. Thus the observed data vector $\phi(t)$ is given by

$$\phi(t) = A(t)\,x(t) \tag{1}$$

where t is time, $A(t)$ is a known matrix, and $x(t)$ is the state of some stochastic process. It is the quantity $x(t)$ which is to be determined. The rank of $A(t)$ is assumed to be less than the dimension of $x(t)$, so that the correct value of $x(t)$ can only be estimated. The estimation procedure to be discussed here will be

[1] Column vectors are small letters; matrices are capital letters; the superscript T denotes the transpose; other superscripts denote scalars; t is time; the star (*) refers to an estimated quantity; the symbol $E\,[\cdots]$ indicates the expected value of the bracketed quantity over the ensemble of all similar experiments.

II. THE SEQUENTIALLY CORRELATED STOCHASTIC PROCESS

The concept of a stochastic process to be considered here follows the discussion in Ref. 1. Briefly, one is to imagine a random process with a time-dependent output vector, $x(t)$, such that each component is a random variable (see Fig. 1). The continuous process can be thought of as the limiting case of a discrete process, which is constructed as follows. Let

$$\{t_0, t_1, \cdots, t_i, \cdots, t_n\}$$

be an arbitrarily chosen set of sample times, and let

$$\{x_0, x_1, \cdots, x_i, \cdots, x_n\}$$

be the corresponding sequence of output vectors on any given experiment, where x_i has been used to denote

$$x_i = x(t_i) \tag{5}$$

If this time sequence is held fixed over the ensemble of all similar experiments there exists a set of time-dependent "first" probability density functions $f^k(x^k, t_i)$, defined over this ensemble, such that

$$\textit{probability}\ (x^k < \xi^k < x^k + dx^k \text{ at } t = t_i) = f^k(x^k, t_i)\, dx^k \tag{6}$$

where x^k is the kth component of the vector x. There also exists a set of time-dependent "second" probability density functions $f^{kl}(x^k, x^l, t_i, t_j)$, such that

$$\textit{probability}\ (x^k < \xi^k < x^k + dx^k \text{ at } t_i,\ x^l < \xi^l < x^l + dx^l \text{ at } t_j) = f^{kl}(x^k, x^l, t_i, t_j)\, dx^k\, dx^l \tag{7}$$

For $t_i = t_j$ Eq. (7) describes the joint probability density function for the random variables x_i^k and x_i^l. Similarly, there exists an nth-order probability density function of the stochastic process. (Only the second-order functions will be needed for the analysis presented here). If it is assumed that the probability density functions introduced above exist for all times t_i, and are continuous functions of t_i and t_j, the stochastic process will be said to be continuous.

for all $t_i \leq t_j \leq t_k$, where $R(t_j, t_i)$ is the normalized correlation matrix defined by Eq. (12).

It is shown in Ref. 2 that a sequentially correlated process is a Markoff process if the x_i^k are Gaussian random variables.

The covariance matrix describing the error in this estimate is

$$\hat{\Lambda}_j = E\,[(\hat{x}_j - x_j)\,(\hat{x}_j - x_j)^T] = \Lambda_j - R_{ji}\,[\Lambda_i - \Lambda_i^*]\,R_{ji}^T \tag{18}$$

Upon introducing the data vector ϕ_j, the minimum variance estimate of x_j becomes

$$x_j^* = \hat{x}_j + W_j\,(\phi_j - A_j\,\hat{x}_j) \tag{19}$$

where

$$W_j = \hat{\Lambda}_j\,A_j^T\,[A_j\,\hat{\Lambda}_j\,A_j^T]^{-1} \tag{20}$$

and the covariance matrix describing the error in this estimate is

$$\Lambda_j^* = E\,[(x_j^* - x_j)\,(x_j^* - x_j)^T] = \hat{\Lambda}_j - W_j\,A_j\,\hat{\Lambda}_j \tag{21}$$

The process is repeated at time $t_k > t_j$, yielding a sequential estimate of the x_i.

Thus, to first order in Δt,

$$[A_j \hat{\Lambda}_j A_j^T] = [A_i Q_i A_i^T] \Delta t \tag{30}$$

$$W_j = [\Lambda_i^* \dot{A}_i^T + Q_i A_i^T] [A_i Q_i A_i^T]^{-1} \tag{31}$$

$$[\Lambda_j^* - \Lambda_i^*] = \{Q_i - W_j [\dot{A}_i \Lambda_i^* + A_i Q_i]\} \Delta t \tag{32}$$

Note that W_j is specified in terms of quantities defined at t_i. (The matrix inverse in Eq. (31) is assumed to exist.) Dividing both sides of Eq. (32) by Δt, and taking the limit as $\Delta t \to 0$, the (nonlinear) differential equation for the covariance matrix of the error in the estimate $x^*(t)$ is found to be

$$\frac{d}{dt} [\Lambda^*] = Q - W [\dot{A} \Lambda^* + A Q] \tag{33}$$

where $Q(t)$ is defined by Eq. (28) and $W(t)$ is defined by Eq. (31). The subscript i has been dropped in Eq. (33), indicating that all quantities are continuously evaluated at the variable time t. Equation (33) can be integrated to obtain $\Lambda^*(t)$.

From Eq. (19) it follows that

$$A_i x_i^* = \phi_i \tag{34}$$

Thus Eq. (19) becomes, to first order in Δt,

$$[x_j^* - x_i^*] = \{\dot{R}_i x_i^* + W_j [\dot{\phi}_i - A_i \dot{R}_i x_i^* - \dot{A}_i x_i^*]\} \Delta t \tag{35}$$

Proceeding as with Eq. (32), the (linear) differential equation for the estimate $x^*(t)$ is found to be

$$\frac{d}{dt} [x^*] = Mx^* + W \dot{\phi} \tag{36}$$

V. AN APPLICATION OF THE ESTIMATION TECHNIQUE

Consider the simple estimation problem presented in Section I, where the observation is

$$\phi(t) = x^1(t) + x^2(t) = A\,x(t) \tag{38}$$

and $A = [1, 1]$. The quantity $x^T(t) = [x^1(t), x^2(t)]$ is the vector to be estimated. Let

$$\Lambda(t) = E\,[x(t)\,x^T(t)] = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} = \text{constant} \tag{39}$$

$$R(t_j, t_i) = E\,[x(t_j)\,x^T(t_i)]\,\Lambda^{-1}(t) = \begin{bmatrix} \exp -\alpha(t_j - t_i) & 0 \\ 0 & 1 \end{bmatrix} \tag{40}$$

and hence

$$\dot{R}(t) = \left[\frac{\partial R(t_j, t_i)}{\partial t_j}\right]_{t_i = t_j = t} = \begin{bmatrix} -\alpha & 0 \\ 0 & 0 \end{bmatrix} \tag{41}$$

Thus the particular process in mind is composed of an exponentially correlated variable $x^1(t)$ with constant variance σ_1^2, and an unknown constant x^2, with variance σ_2^2. The cross correlations are zero. The estimation problem could be interpreted as the extraction of the constant x^2 from the "noise" $x^1(t)$, or as the recovery of the "signal" $x^1(t)$ in the presence of the measurement bias x^2.

It will be convenient to define the matrix

$$\Delta(t) = \Lambda - \Lambda^*(t) \tag{42}$$

where $f(t) = 2\,(r^2 + \alpha t/2 + 1)$. Equation (48) integrates to yield

$$x^{1*}(t) = \phi(t)\left(\frac{f(t)-1}{f(t)}\right) - \phi(0)\left(\frac{2r^2+1}{f(t)}\right) - \left(\frac{\alpha}{f(t)}\right)\int_0^t \phi(\tau)\,d\tau \tag{49}$$

$$x^{2*}(t) = \phi(t) - x^{1*}(t) \tag{50}$$

Notice that as $\alpha \to \infty$ and/or as $t \to \infty$,

$$x^{2*}(t) \to \frac{1}{t}\int_0^t \phi(\tau)\,d\tau \tag{51}$$

VII. DISCUSSION OF THE RESULT

It has been shown that the estimation equation for a discrete, sequentially correlated stochastic process extends to the continuous case to yield a linear differential equation for the time-varying estimate. Only the *rate of change* of the observed data vector enters into this equation, which is a consequence of applying the relationship (34). The reader may be tempted to carry this process one step further, and substitute

$$\dot{\phi}(t) = \dot{A}(t)\, x^*(t) + A(t)\, \dot{x}^*(t) \tag{56}$$

into Eq. (36), thereby eliminating any consideration of the data at all. It can be quickly verified, however, that this leads to a differential equation for the estimate which has no unique solution.

In order to solve Eqs. (33) and (36) numerically with a digital computer it may be convenient to express them as difference equations; i.e.,

$$\Lambda^*_{i+1} = \Lambda^*_i + \left(Q_i - W_{i+1}\, [\dot{A}_i\, \Lambda^*_i + A_i\, Q_i]\right) (t_{i+1} - t_i) \tag{57}$$

$$x^*_{i+1} = (R_{i+1,i})\, x^*_i + W_{i+1}\, [\phi_{i+1} - (A_{i+1}\, R_{i+1,i})\, x^*_i] \tag{58}$$

where Q_i and W_i are evaluated from Eqs. (28) and (31). Equations (57) and (58) are the limiting cases of (19) and (21) as $(t_{i+1} - t_i) \to 0$. It is interesting to note that in passing from the discrete to the continuous case certain new matrices appear, such as $\dot{\Lambda}$, $\dot{A}$, and $\dot{R}$, and other combinations of terms go to zero. The difference equations (57) and (58) could therefore experience different numerical behavior than Eqs. (19) and (21) when being evaluated on a digital computer, even though the results are theoretically identical in the limiting case $\Delta t \to 0$. Thus the continuous form of the estimate might be preferable for some practical problems where numerical stability must be considered.

REFERENCES

1. Reza, F. M., *An Introduction to Information Theory,* ch. 10, McGraw-Hill Book Co., New York, 1961.

2. Pfeiffer, C. G., *Sequential Estimation of Correlated Stochastic Variables,* Technical Report 32-445, Jet Propulsion Laboratory, Pasadena, July 1, 1963.